DOCUMENT RESUME

ED 275 722 TM 860 636

AUTHOR Centra, John A.
TITLE Colleague Evaluations: The Critical Link.
PUB DATE Apr 86
NOTE 6p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Viewpoints (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *College Faculty; *Evaluation Methods; Evaluation Problems; Higher Education; Holistic Approach; Observation; *Peer Evaluation; *Teacher Evaluation; Weighted Scores

ABSTRACT

Although college faculty may hesitate to evaluate each others' teaching, it is a sensible alternative to evaluation by students or by administrators. Peer evaluations are not only useful for improving teaching, but also affect promotion and tenure decisions. When faculty are rated by observation, repeated visits and rater training are required in order to insure reliability. A qualitative approach, involving descriptions of classroom instruction based on the observer's perceptions, is preferred. Rating scales may be useful in assessing the appropriateness of course objectives, value of instructional materials, student achievement, and teacher's knowledge of subject matter. One college uses a Committee on Teaching to evaluate teaching performance. In addition to reviewing materials presented by the instructor and describing classroom observations, the committee interviews both former and current students. Another approach to teacher evaluation, which is used by very few institutions, is a variable weight approach. Faculty members select, in advance, an area of their performance--research, teaching, or service--which would be given extra weight during that evaluation year. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED275722

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. A. Centra
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Colleague Evaluations: The Critical Link

John A. Centra[1]
Syracuse University

Faculty members have little hesitancy about asking a valued colleague for reactions to a draft research article or proposal, but much less frequently do they ask a colleague to react to a course syllabus or to visit their class and make suggestions about their teaching. Why? Is it that faculty believe teaching is more personal and subjective, whereas the standards of good research or scholarship are more widely known and objective? Is it because visits to classes can be more time consuming and are potentially sensitive? Is it because research is more important in most settings? Or perhaps it's because one's scholarship, if published, is there for all to see and evaluate, while teaching is seen "only" by one's students.

My guess is that it's some of all of the above. And unless faculty members are willing to leave the evaluation of teaching to students, who possess only a limited view, or to administrators, who often don't have the time or the necessary background, then they must be willing to invest their time and effort in peer evaluation of teaching as well as of research. While these evaluations can be useful for improving the level of teaching -- formative evaluation -- they may be even more

---

[1] Paper given at the 1986 American Educational Research Asssociation Meeting, San Francisco, CA.

TM 860 636

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

critical in affecting promotion and tenure decisions for it is that process that expresses the values of an institution and it is that process that determines who one's colleagues will be for years to come.

Some people view peer evaluation of teaching as synonomous with classroom observations. Observations, however, are only one facet of the evaluation and not a particularly reliable one at that. Research has shown that when colleague ratings are based solely on what they observe in the classroom there is only slight interrater agreement (Centra, 1975). This same research also demonstrated that the ratings tend to be very generous, even when they are not shared with the teacher rated. Reliable classroom observations probably require that faculty members participate in some type of training to establish a common basis for making the quantitative ratings. It's unlikely that many faculty members will want to take the time for such training or to make the many repeated visits that would provide a satisfactory sampling of teaching behavior. A different approach is therefore needed, one perhaps that uses qualitative rather quantitative principles of evaluation. A qualitative approach would involve descriptions of classroom instruction based on the perceptions of the observers. Rating scales and numerical judgments would not be necessary or useful. In many instances, in fact, the scales are not appropriate to all styles of teaching. But descriptions by several observers will more likely reflect possible biases and

the resulting narrative could be much more useful for summative or formative purposes.

Teaching involves much more than what goes on in the classroom and colleagues should review these aspects of teaching as well. Included are such aspects as the level and appropriateness of objectives in the course, the design and value of instructional materials, the level of student achievement, and the teacher's knowledge of subject matter. Although peer assessments of these criteria could be based on such tangible items as the course syllabus and examination questions and results, there is at this time no evidence that these assessments will be valid or reliable. Rating scales and quantitative judgments should be more useful here than with classroom observations but that remains to be demonstrated. The same basic problem may exist: lack of common standards among untrained ratings.

In addition to what and how colleagues might best judge teaching, there is also the issue of which colleagues should be involved when tenure or promotion decisions are being made. Tenure and promotion committees always include faculty representatives but generally they don't have the time or manpower to obtain their own independent information on teaching effectiveness. Some institutions form an ad hoc committee for each faculty member being considered and these committees typically consider all of the evidence on teaching, research and service and then make a recommendation to the tenure and

promotion committee. These ad hoc committees may consist of an anonymous group of colleagues, a committee that is selected randomly or purposely from an appropriate sub-group of faculty, or other variations (Centra, 1979). One college uses an approach that I think could be a useful model for others to consider. At this college, an ad hoc committee is appointed for each faculty member up for tenure or promotion and this group, called a Committee on Teaching, looks only at teaching performance. The committee consists of a senior faculty member, a junior faculty member, and a student, all from outside the candidate's department. The candidate provides the committee with a variety of teaching-related material: copies of syllabi, exams, project reports, student evaluations, teaching aids, and the like. The candidate also draws up a list of student advisees, dissertation advisees and names of past and current students. Through phone calls and letters, the committee is able to piece together a picture of student views of the teacher to supplement the student evaluation information provided by the teacher.

Members of the committee also observe at least one class, and write up a description of what they saw. The Committee on Teaching's report to the tenure and promotion committee then includes descriptive information and evaluations based on their observations, review of materials provided, and conversations or letters with students and advisees. Candidates typically cooperate fully with their committee, for to do otherwise would in itself reflect their disinterest in teaching. Interesting

enough, this college is part of a research-oriented university where good teaching alone is not sufficient; but at least there is a systematic effort to include teaching effectiveness in summative deliberations. More importantly, the effort goes beyond student ratings by requiring colleagues to look at a variety of evidence.

Many institutions may give more weight to research and scholarship because it is more easily quantified and because the variability among faculty members is greater than it is for teaching. Colleague evaluations, through a procedure such as the Committee on Teaching approach, may serve to provide some balance to the process. Another practice, used by very few institutions to my knowledge, is a "variable weight approach." This allows faculty members to select, a year in advance, an area they would like to have receive additional emphasis in performance evaluation. Thus, a minimum of 35 percent of an individual's evaluation might be in research, another 35 percent in teaching, and 10 percent in service. The remaining 20 percent could be allocated to any of the three areas depending on an individual's activities in any given year. Extra time and effort in teaching in any given year, for example, would be given additional weight and depending on the quality of this effort, might result in a more positive evaluation of the faculty member.